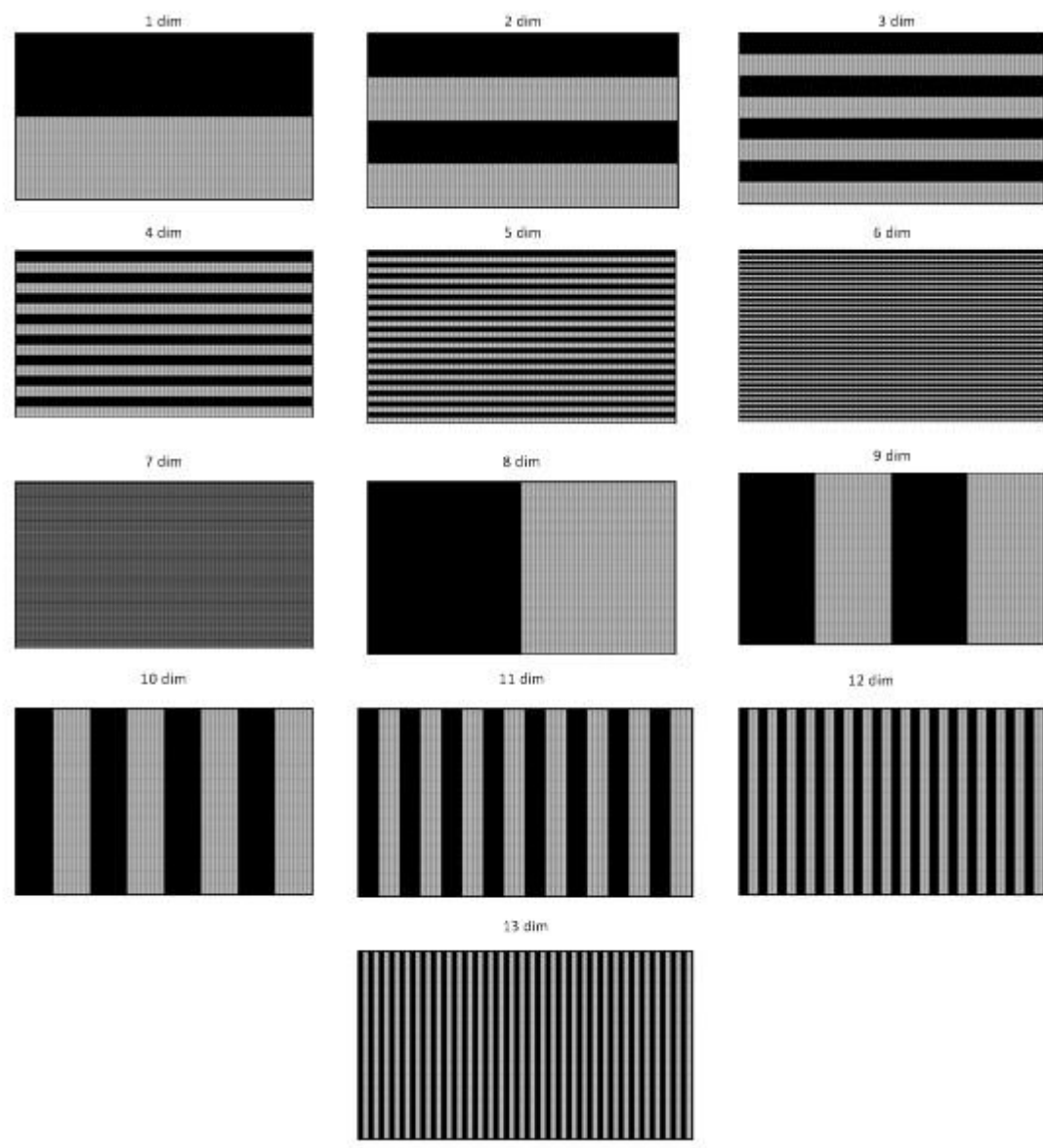


## 本周工作

1. 办理淘宝离职手续；
2. 标签关系研究，结果如下：

图一：13 个单维分布图。可以看到维度呈现先按 Y 轴细分，再按 X 轴细分的趋势；从图中只能看到两个方向上的对称性，第一维和第八维出现了明显的可辨识特征，其他维度则相对困难。有没有办法做到十三个方向的对称且保证分块大小一致呢？



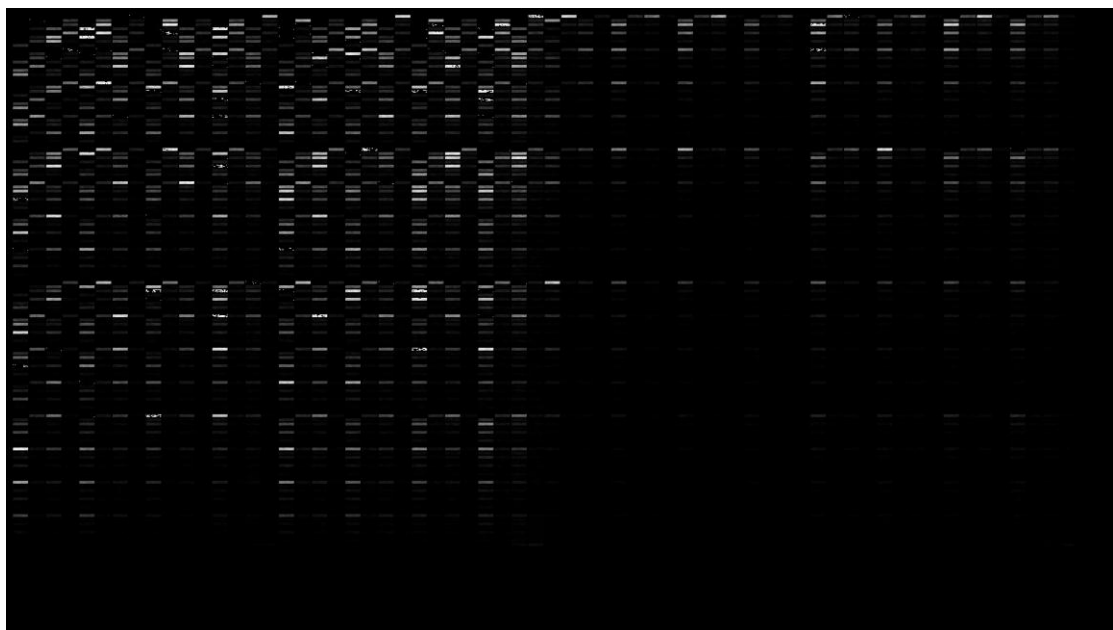
图二：所有数据的 KDE 结果图

作法：

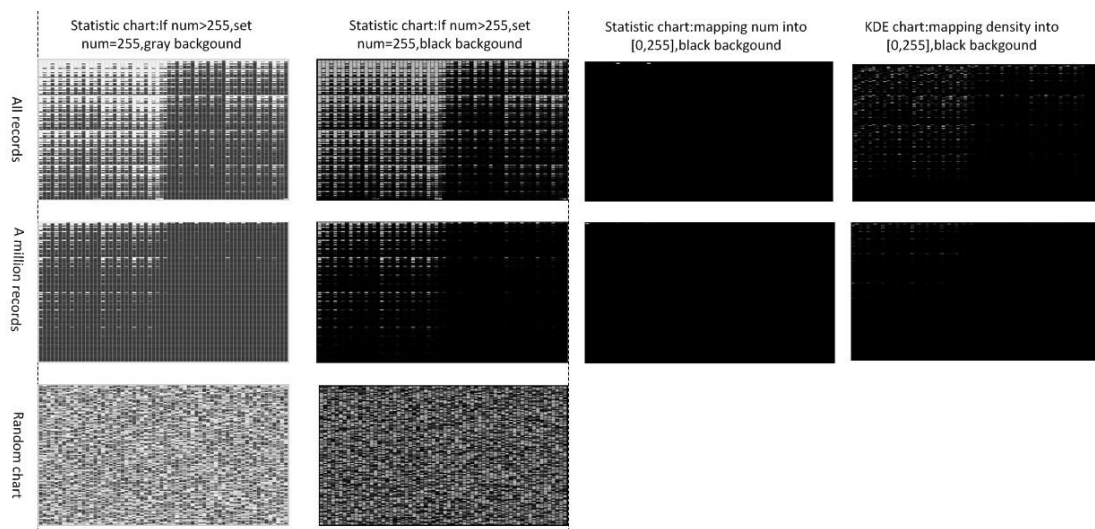
1. 按照每个分类小块的像素个数确定 KDE 的值域，比方说  $50 \times 20 = 1000$ ；
2. 将数据均分成 1000 份，在每份中统计属于 8192 个分类的个数，放在一个二维数组中：  
`int kdeSample[8192][1000];`

3. 按照上面得到的数组计算 KDE（如下），并将得到的  $\text{density} \times 255$  转化为灰度值。

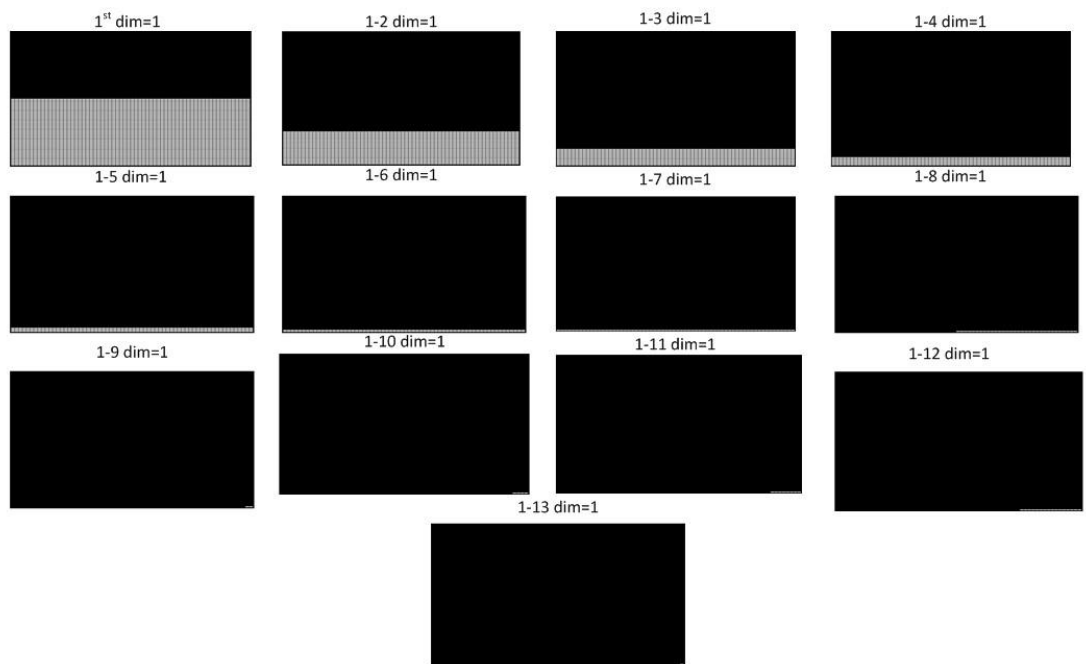
```
double h = 0;
for (int i = 0; i < categorySize; i++) {
    for (int j = 0; j < kdeUnitPointsNum; j++) {
        double result = 0;
        for (int k = 0; k < kdeUnitPointsNum; k++) {
            result += kdeSample[i][k] * Math.exp(-Math.pow((double)(k-j)/h, 2) / 2);
        }
        result = result / (kdeUnitPointsNum * h * Math.sqrt(2 * Math.PI));
        kdePoints[i][j].setColor((int)(255 * result));
    }
}
```



图三：统计与 KDE 对比图。可以看到 KDE 图和统计图整体上一致，但也有不一样的地方，如 KDE 的亮度以及亮块个数都没有统计图多。如何解释，我还在考虑。



图四：多维度关系图。按从左到右、从上到下的顺序，依次是不同维度交错的结果图；由于是按数字顺序依次增加维度，结果上也呈现规律性递减的趋势；怎么达到快速识别多个维度仍是一个需要认真研究的问题，而这与图一中的问题有非常紧密的连续，个人觉得这也是像素创新的关键。



## 下周计划

1. 背完托福单词（第一遍）；
2. 明确上述问题的关键和努力方向，继续研究。